University of Huddersfield Repository

Lee, Hyunkook and Rumsey, Francis

Investigation into the Effect of Interchannel Crosstalk in Multichannel Microphone Technique

Original Citation


This version is available at http://eprints.hud.ac.uk/id/eprint/9682/

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

http://eprints.hud.ac.uk/
Investigation into the Effect of Interchannel Crosstalk in Multichannel Microphone Technique

Hyun-Kook Lee and Francis Rumsey

Institute of Sound Recording, University of Surrey, Guildford, Surrey, GU2 7XH, England
Correspondence should be addressed to Hyun-Kook Lee: tonmeister1025@yahoo.co.uk

ABSTRACT

A series of subjective listening tests were carried out in order to investigate the effect of interchannel crosstalk in multichannel microphone technique. Perceived attributes of interchannel crosstalk images were first elicited, and then graded with various independent variables, including different types of microphone array (different combinations of time and intensity differences), sound source and acoustic condition. The results showed that the most dominant effects of interchannel crosstalk were an increase of source width and a decrease of locatedness. The ratio of time and intensity differences in microphone array was the most significant factor for both effects. Sound source type had a significant effect for source width increase, but not for the locatedness decrease. Acoustic condition was significant for locatedness decrease, but not for source width increase. This paper describes the experiment method, and presents and discusses the details of the result data.

1. INTRODUCTION

This paper examines subjective effects of interchannel crosstalk in multichannel microphone techniques for classical music recordings. Since multichannel stereophonic audio systems such as the so-called 5.1 surround became popular in recent years, a number of multichannel microphone techniques have been proposed corresponding to the requirement of the new system. Multichannel stereophony is able to overcome some of the limitations of conventional two-channel stereophony, by adding a centre channel providing a stable centre image and two surround channels delivering a sense of spatial impression. However, the addition of extra channels inevitably gives rise to a question about the effect of interchannel crosstalk. If a three-channel microphone technique is used for recording a single sound source located at a certain position, each pair of microphones (L-C, R-C and L-R) will pick up the sound with a different relationship of time and intensity differences.

The influence of interchannel crosstalk on the resulting image quality has been an issue for debate. When
talking about interchannel crosstalk, there is the implicit assumption that signals from microphones other than the pair primarily covering the sector of the recording angle in which a source lies can be treated as unwanted ‘crosstalk’. This suggests that two-channel stereophonic principles are being applied to the analysis of multichannel microphone techniques, which is, by and large, the principle on which the arrays discussed in this paper have been designed. Theile [1] asserted that the ideal localisation curve for a three channel microphone array should be linear within the recording angle, similarly to that for a conventional two channel array. In order to achieve this ideal, the sound source located in one side of the recording angle must not be picked up by the microphone in the other side. Although this is practically impossible, Theile [1] suggested that a reasonably balanced localisation could be achieved by reducing the intensity of the interchannel crosstalk signal as much as possible. His claim is that if the interchannel crosstalk is not suppressed enough, triple phantom images will be created in reproduction, thus leading to a decrease in image focus and clarity. Intending to optimise the operation of three channel microphone technique with respect to interchannel crosstalk, a microphone technique called OCT (Optimised Cardioid Technique) has been proposed. The details of this technique are found in [2].

Theile’s hypothesis of the perception of three separate images was claimed in [3]. Rumsey asserted that the listener tends to perceive a single fused phantom source whose ‘size, stability and position are governed by the relevant intensity and time differences between the signals’, and suggested the need for further experiments regarding the effect of interchannel crosstalk. In fact, there is no experimental evidence available to support the triple phantom image hypothesis.

Williams [4] disagrees about the significant effect of interchannel crosstalk that was claimed by Theile. His argument is that the interchannel crosstalk is already reduced to a great extent using directional microphones, and therefore is not a great matter for localisation. He seems to suggest that it is more important to link the recording angles of each stereo segment without overlap in order to obtain a balanced localisation performance, rather than achieving the maximum suppression of interchannel crosstalk. Williams and Le Du [5] proposed a microphone technique based on their so-called ‘critical linking’ concept.

Although there is much debate on the issue of so-called interchannel crosstalk as shown above, to date there is no conclusive answer as to whether it really matters or not. More importantly, there seems to be no research of which the authors are aware that investigates the subjective effects of interchannel crosstalk in such contexts. Therefore, it is not clearly known what kinds of interchannel crosstalk effects listeners perceive and how those factors affect the stereophonic image quality. This is why the current experiment was designed. It is expected that the results of this experiment will provide recording engineers with useful guidelines for the design and application of multichannel microphone techniques. The primary research questions for this experiment were formulated as follows:

- What subjective attributes are perceived as a result of interchannel crosstalk?
- How audible are these attributes in general?
- Does the subjective grading for these attributes depend on type of microphone array (combination ratio of time and intensity differences), type of sound source, or acoustic condition?
- Is there a significant interaction between each of these variables?

2. EXPERIMENTAL DESIGN

2.1. General Experimental Methodology

The basic concept for the method used in the current experiment was inspired by a test method known as QDA (Quantitative Descriptive Analysis). A basic QDA consists of three stages [6]. Firstly, a group of qualified subjects are presented with stimuli and develop descriptive terms on the attributes of the product through discussion. Secondly, the elicited terms are grouped into common attribute scales through discussion based on the similarity of meaning of the terms. Finally, the stimuli are graded using the obtained common attribute scales. The advantages of using this method would be firstly that any experimenter’s bias on the scales to be graded can be avoided, and secondly that the data obtained from a number of subjects can be statistically analysed together. However, it was considered that a full QDA, according to methods proposed in the food sciences, would be an unnecessarily time consuming and detailed process that
would require gathering all subjects at the same time and place for a series of group discussions. Therefore, the QDA method was modified in several ways so as to select only the parts relevant to the problem in hand and to adapt it to the current experimental context in sound recording. Firstly, instead of using a full elicitation process, subjects were asked to select relevant attributes from a set of potential attributes that were provided. They were also asked to describe additional attributes using their own terms in case they perceived any other differences than the ones provided. The elicited subjective terms were carefully interpreted and unified by the experimenter through later informal discussions with individual subjects on the meanings of the terms they used. Subjects were also asked to grade the magnitude of audibility of the selected attributes. This was in order to weight the perceptual dominance of those attributes and accordingly reduce the number of attribute scales to be graded. Finally, a grading test was conducted using the selected attribute scales and statistical tests were carried out on the obtained data.

From the above descriptions, it can be said that the whole experiment consisted of two stages: attribute selection test and grading test. More detailed descriptions of the test methods are presented in sections 3.1 and 4.1.

2.2. Choice of Microphone Technique

2.2.1. Basic philosophy

To date, a number of microphone techniques have been proposed for the recording and reproduction of multichannel surround sound. Rumsey [3] suggested a way of classifying the design concepts of the current multichannel microphone techniques, based on the purpose of the rear channels in the technique. According to his classification, there are two main groups: those that use a ‘five-channel main microphone technique’ and those that use a ‘technique with front and rear separation’. The former consists of five microphones that are placed relatively close to each other and form a single array, pursuing the recreation of a natural sound field of the recording space. In other words, the microphone techniques in this group attempt to provide a satisfying directional image and spatial impression at the same time, with a fixed pattern of microphone placement. The fixed positions and polar patterns of the front and surround microphones might result in an inevitable compromise between the representations of optimised directional image and spatial impression. For example, the front triplet should be optimised not only with respect to the recording angle of direct sound from the front but also with respect to the balance of direct and indirect sound intensity in conjunction with the surround microphones [2]. The position and the polar pattern of the surround microphone array should not only be decided for the characteristics of the ambient sound, but also for the suppression of the direct sound due to the relatively short distance between the front and the rear microphones. With this technique, interchannel crosstalk will be an issue not only between the front channels but also between the front and surround channels due to the relatively short distance between the front and surround microphones.

The other group, on the other hand, uses a frontal main microphone array that is used specifically for accurate pickup of direct sound so that sources can be easily localized on reproduction, together with a separate rear microphone array that is designed to pick up decorrelated ambient sound to feed the surround loudspeakers. Different rear microphone arrays can be combined with different frontal arrays depending on desired directional and ambience characteristics [2]. The distance between the front and the rear arrays can vary depending on different recording situations. Interchannel crosstalk between the front and rear microphones would not be taken into account much with this type of technique because of the sufficiently long distance between them. In this regard, it seems that this type of technique gives recording engineers more freedom to control the spatial impression and enables them to use their artistic and technical creativity more than the 5-channel main microphone technique. For this reason, a technique with separate treatment of front and rear was chosen as the basis for this experiment.

2.2.2. Simulation of microphone technique

It was considered that if a microphone technique were operated in a practical recording venue, such uncontrolled acoustic artefacts as reflections and reverberation might lead to difficulty when analysing the factors that caused the resulting perceptual effects. In order to obtain data about the effects of interchannel
crosstalk on phantom images in the absence of room reflections the experiment included a simulation of recordings made in an anechoic condition, rather than using recordings made in a practical venue. For the anechoic experiment, only a 3-channel frontal microphone technique was needed. Even though understanding the effect of interchannel crosstalk in anechoic recording conditions was the primary aim of this research, it was also of interest to see how the perception of this effect would differ in the context of different reverberant recording conditions. As discussed in the previous section, the purpose of rear microphone array in the context of this experiment is to provide a diffuse ambience rather than a localisable image of the direct sound. The ambience picked up by a rear microphone array was simulated by using an artificial reverberator.

2.2.3. Frontal microphone technique

The frontal microphone technique chosen for this experiment was the so-called ‘critical linking’ 3 channel microphone technique, proposed by Williams and Le Du (detailed descriptions of this technique can be found in [5]). The basic design concept of this technique aims to achieve a continuous distribution of phantom images across channels L-C-R by linking the recording angles of each stereophonic segment C-L and C-R without overlap. Within one segment, the psychoacoustic laws for localisation in conventional 2 channel stereophonic reproduction such as ‘summing localisation’ or the precedence effect are applied independently without considering the influence of the other segment. For example, when a sound source is located at 45 degrees to the right of the centre line, localisation of the phantom image should be governed by the summing localisation effect between C and R only, and in this case L should not be regarded as a crosstalk to the channels C and R. Ideally, L should not be taken into account in localisation process since it is to be suppressed by the same effect or the precedence effect operating between C and L. It is hypothesised that even though the position of the phantom image can be solely determined by C-R without the aid of L, the presence of L will influence the spatial or timbral quality of the image to some extent. Reported studies on the perceived differences between phantom images created by the precedence effect and their corresponding mono images could be the basis for this hypothesis (e.g. the phantom image having ‘greater spatial extent’ [7], ‘image extended toward the echo source’ [8] and ‘fuller tonal colour’ [9]). In this regard, it is logical to examine the effect of interchannel crosstalk by comparing the image that is created with the crosstalk channel turned on (image formed by contributions from L-C-R) and that with the crosstalk channel turned off (C-R only).

The critical linking technique supposedly enables one to create various array styles having different distances and angles between microphones while keeping the recording angle across L-C-R constant. Therefore, the effect of the ratio between time and intensity differences between the crosstalk and the other channels can be investigated by comparing different microphone arrays sharing the same recording angle. Williams provided various examples of critically linked microphone arrays that can be created while keeping the same recording angle. For the current experiment, only four representative arrays were selected from the examples, as shown in Figure 2.1 to Figure 2.4. The recording angles for these arrays were all 180°. The simulated direction of the sound source was 45 degrees from the centre line of the array, and the distance from the centre point of the array was 5 metres. The interchannel time and intensity differences between L and C and those between R and C calculated for each array according to these are shown in Table 2.1. In a conventional 2 channel stereophonic reproduction, the minimum delay time required between left and right channels for operating the precedence effect is normally 1.1ms for natural sound sources ([10], [11] and [12]). According to Theile’s hypothesis [2], in a three-channel reproduction, the same effect between centre and left (or right) channels can be achieved with only a half of this delay time (0.55ms). From this, it might be possible to assume that the delay times between C and L in all the arrays shown in Table 2.1 are long enough to operate the precedence effect between C and L.

<table>
<thead>
<tr>
<th>Array</th>
<th>C to L delay</th>
<th>C to L intensity</th>
<th>C to R delay</th>
<th>C to R intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.64ms</td>
<td>-20.5dB</td>
<td>0.08ms</td>
<td>-0.7dB</td>
</tr>
<tr>
<td>2</td>
<td>0.79ms</td>
<td>-12.8dB</td>
<td>0.06ms</td>
<td>0.6dB</td>
</tr>
<tr>
<td>3</td>
<td>0.94ms</td>
<td>-8.0dB</td>
<td>0.16ms</td>
<td>1.2dB</td>
</tr>
<tr>
<td>4</td>
<td>1.09ms</td>
<td>-4.6dB</td>
<td>0.21ms</td>
<td>1.4dB</td>
</tr>
</tbody>
</table>

Table 2.1: Time and intensity differences of the left and right channels to the centre channel for each array.
2.3. Choice of Sound Source

It was of interest to examine whether the effect of interchannel crosstalk depends on the use of different types of sound source. Three types of natural sound sources comprising cello, bongo and speech were chosen for this experiment due to their distinctive temporal and spectral characteristics, with the cello being relatively continuous and having a complex harmonic structure, the bongo having a strong transient nature, and the speech having a fine mixture of transient and continuous sounds as well as a wide range of frequencies. The signal for each sound source was an anechoic mono recording of a performance excerpt taken from the Bang & Olufsen Archimedes project CD [13]. From a psychophysical viewpoint, it might be claimed that the characteristics of natural sound source are too complex to strictly analyse the effect of spectral or temporal characteristics of the sound. In fact, the use of pure sine tones or bandpass noise signals might allow a more controlled investigation on various aspects. However, the results obtained with strictly controlled stimuli might not be applicable to natural sound sources in the same manner because the characteristics of the latter are more complex, and therefore it was deemed to be more appropriate to use sound sources likely to be encountered in practical recording situations. The waveform and frequency analysis plot for each sound source are shown in Figures 2.5 and 2.6. The
waveform shows temporal variations during specific 0.3 seconds taken from the whole performance, and the frequency analysis is a plot of the average intensity by frequency over the whole performance.

Figure 2.5: Extracts of waveforms of each sound source used for the experiment

(a) Cello source

(b) Bongo source

(c) Speech source

Figure 2.6: Long-term averaged frequency spectrum of each sound source used for the experiment

2.4. Acoustic Conditions

The acoustic conditions considered for this experiment comprised anechoic, room and hall. As mentioned above, the anechoic condition was of primary interest since it enabled the strict control of variables, and it was created naturally by using anechoically recorded sounds sources. Simulations of recordings made in different acoustic conditions were also used in order to predict the behaviour of interchannel crosstalk in practical recording venues such as room and hall. The room and hall conditions are chosen for their different acoustical characteristics. The detailed characteristics of the simulated room and hall conditions are described in the next section.
2.5. Stimuli Creation Process

A set of multichannel stimuli, involving 36 combinations of four microphone arrays, three sound sources and three acoustic conditions, was processed for the experiment. The process was carried out in Studio 3, a multichannel sound control room of the University of Surrey’s Department of Music and Sound Recording. The diagram for the stimuli creation process is shown in Figure 2.7. For the creation of the anechoic stimuli, monophonic signals of each anechoic sound source were first fed into three separate channels on a Sony Oxford-R3 digital console, and they were processed in accordance with the time and intensity relationship of each microphone array shown in Table 2.1. The processed signal of each channel was then routed to each group output of L, C and R for the reproduction of three front channels. On the other hand, the room and hall stimuli were mixed for the reproduction of all five channels. The monophonic signal of the anechoic sound was sent to a Lexicon 480L reverberator through an auxiliary output of the mixer. The four purely ambient output signals generated from the reverberator were then routed to two group outputs for reproduction of the front channels L and R as well as those for the surround channels LS and RS, with the intensities of each signal kept the same, thus being mixed with the original anechoic sound signals in L and R. The basis for using the four outer channels for reproduction of the reverberation signals is as follows. Hiyama et al [14] investigated the number of loudspeakers required for the reproduction of optimum spatial impression of diffuse sound field. A reference loudspeaker arrangement consisting of 24 loudspeakers placed at every 15° making a circle was compared with various arrangements having a different number of loudspeakers (12, 8, 6, 5, 4, 3 and 2) with regard to spatial impression. They found that at least 4 loudspeakers, which were arranged in similar positions as the BS.775-1 recommendation, were required for listeners to perceive a similar spatial impression to the reference sound. For creating ambience sounds of room and hall, the presets of ‘large room’ and ‘large hall’ setup existing in the reverberator were used. The details of the reverberator set up used for creating the room and hall ambience sounds are shown in Table 2.2. In general, the ‘large room’ set can be described as producing coloured and comb-filtered ambience sounds with slapping echoes. The ‘large hall’ creates ambience sounds that have longer reverberation time and are more diffused without colouring the direct sound.

<table>
<thead>
<tr>
<th></th>
<th>Size</th>
<th>RT Mid</th>
<th>RT Low</th>
<th>HF Cut-off</th>
<th>Pre-delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large Room</td>
<td>19m²</td>
<td>0.70s</td>
<td>0.70s</td>
<td>6.593kHz</td>
<td>0ms</td>
</tr>
<tr>
<td>Large Hall</td>
<td>37m²</td>
<td>2.19s</td>
<td>2.63s</td>
<td>2.862kHz</td>
<td>24ms</td>
</tr>
</tbody>
</table>

Table 2.2: Parameters of the reverberation setup used for simulations of room and hall (RT Mid = middle frequency reverb. time, RT Low = low frequency reverb. time)

The mixing ratio of the direct sound and reverberation was up to the authors’ aesthetic judgement as experienced balance engineers, and it was aimed to compromise between maintaining the clarity of the direct sound and achieving sufficient listener envelopment. The signals from each group output were individually recorded to computer hard disk using a Protools hard disk recording interface, and were eventually transformed as monophonic audio files.

![Figure 2.7: Diagram of signal processing for stimuli creation](image-url)
2.6. Physical Setup

The experiment was conducted in an ITU-R BS.1116-compliant [15] listening room at the University of Surrey. According to the ITU-R BS.775-1 recommendation [16], five Genelec 1032A loudspeakers were set up at 0°, 30° and 100°, with a distance of 2m from the subject’s seat. The levels of the loudspeakers were aligned to be equal, and the sound pressure level of all stimuli was calibrated at 75dB, A weighted, at the listening position. The stimuli were played back through a Yamaha O2R mixing console, and controlled by a computer-based control interface placed in front of the listener’s seat.

2.7. Test Subjects

A total of eight subjects took part in the experiment. All were experienced listeners, selected from staff members, research students and final year undergraduate students on the University of Surrey’s Tonmeister course.

3. ATTRIBUTE SELECTION PROCESS

3.1. Method

This process used only six representative stimuli from the whole set of stimuli created. They were each anechoic sound source combined with microphone arrays 1 and 4, which were considered to have the most distinctive difference in perception of the resulting images. The reason for using the anechoic stimuli only was that they were considered to enable the most focused listening to the effect of interchannel crosstalk without any artefacts of recording room acoustics. This test was designed to give the subject the freedom to control the playback of the stimuli. Figure 3.1 shows the control interface used for this test, which was written using the MAX-MSP software. There were a total of six trial pages, and the buttons A and B in each page presented the images of C-R and L-C-R in random orders. The stimuli pair of A and B was synchronised and looped so that the subjects were able to switch between them freely, and to listen repeatedly.

Figure 3.1: Layout of the control interface used for the pairwise comparison and elicitation of auditory attributes

There were two tasks for the subjects to complete in this test, comprising:

1. To define the global set of auditory attributes for the perceived differences between the images of C-R and L-C-R
2. To grade the overall intensities of audibility for those attributes

The first task was given in order to understand the basic auditory percepts arising from interchannel crosstalk. As discussed in section 1.1, the subjects were provided with a list of potential attributes and asked to select the ones relevant to the perceived differences. Any additional differences perceived were also to be described using the subjects’ own terms, and they were to be unified into the common terms by informal discussions between the subjects. The choice of the provided attributes was based on the results of the authors’ previous experiment conducted to investigate the perceptible differences between monophonic and 2-channel stereophonic images [17]. In that experiment, a group of common attributes describing the perceived differences were elicited from subjects, and the effects of interchannel time difference, interchannel intensity difference and the type of sound source on the magnitude of those attributes were examined. The elicited attributes were three spatial and three timbral ones, comprising source width, source focus, source distance, brightness, hardness and fullness. A number of other spatial or timbral attributes are available to
choose from various elicitation experiments ([18], [19] and [20]). However, due to the similarity of the experimental contexts, the attributes perceived between monophonic and 2-channel stereophonic attributes were considered to be the most appropriate basis for evaluating the differences between 2-channel (C-R) and 3-channel (L-C-R) stereophonic images. The choice of attributes and their definitions are shown in Table 3.1. For the attribute meaning the ease of localisation, the term ‘source focus’ from the result of the previous 2-channel experiment was replaced with ‘locatedness’ [21] since the semantic meaning of the former could well be confused with that of ‘source width’. The ‘source location’ attribute was additionally included because a small degree of source location shift was noticed between the images of C-R and L-C-R in the authors’ own informal test.

The purpose of the second task was to limit the number of attributes to be graded in the next test. Grading all the elicited attributes was considered to be ineffective since minor attributes are likely to have small experimental effects. The 10-point scale shown in Figure 3.2 was used for the subjects to grade the magnitudes of audibility of the elicited attributes. The magnitude of audibility might vary for different stimuli, but the grading was to be made to the one having the greatest magnitude.

### Table 3.1: Definitions of the auditory attributes provided to subjects for elicitation

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Source width</strong></td>
<td>The perceived width of a sound source itself&lt;br&gt;i.e. Is one source perceived to be wider than the other?</td>
</tr>
<tr>
<td><strong>Source distance</strong></td>
<td>The perceived distance from the listener to a sound source&lt;br&gt;i.e. Can the sources be discriminated in terms of their distances?</td>
</tr>
<tr>
<td><strong>Source location</strong></td>
<td>The perceived location of a sound source&lt;br&gt;i.e. Does the apparent location of the source appear to change?</td>
</tr>
<tr>
<td><strong>Locatedness</strong></td>
<td>The easiness of localisation of a sound source&lt;br&gt;i.e. How easy is it to pinpoint the apparent location of a source?</td>
</tr>
<tr>
<td><strong>Brightness</strong></td>
<td>The timbral characteristics of a sound depending on the level of high frequencies i.e. bright / dull</td>
</tr>
<tr>
<td><strong>Hardness</strong></td>
<td>The timbral characteristics of a sound depending on the level of mid-high frequencies (typically in the range of 2 – 4kHz) i.e. hard / soft</td>
</tr>
<tr>
<td><strong>Fullness</strong></td>
<td>The timbral characteristics of a sound depending on the level of low frequencies i.e. full / thin</td>
</tr>
</tbody>
</table>

![Figure 3.2: Scale used for grading the audibility of each attribute elicited](image)

**3.2. Results and Discussions**

As a result of the elicitation test, a total of eleven attributes were elicited from the subjects, comprising all of the seven provided attributes and four additional attributes. Table 3.2 shows the attributes elicited, the number of their occurrences, and their audibility indexes. The audibility index represents the average magnitude of audibility for each attribute, and it was obtained by dividing the sum of the audibility grading values obtained for each attribute by the number of subjects. According to the results shown in the table, ‘source width’ appears to be the most audible attribute, having an audibility index of 6.5. The second most audible attribute is shown to be ‘locatedness’. The audibility index is 4.7, and this value indicates that the attribute was audible more than ‘slightly’ according to the semantic labels on the scale. The audibility indexes of all other attributes are shown to be lower than 4.0.
This means that the differences for those attributes were in the range between just audible and slightly audible, which are considered to be minor effects. Therefore, the ‘source width’ and ‘locatedness’ attributes, which were graded above ‘slightly audible’ range, were finally selected to be used for the next grading test.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Occurrences</th>
<th>Audibility index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source width</td>
<td>7</td>
<td>6.5</td>
</tr>
<tr>
<td>Locatedness</td>
<td>6</td>
<td>4.7</td>
</tr>
<tr>
<td>Source location</td>
<td>6</td>
<td>3.6</td>
</tr>
<tr>
<td>Fullness</td>
<td>5</td>
<td>3.5</td>
</tr>
<tr>
<td>Source distance</td>
<td>7</td>
<td>3.1</td>
</tr>
<tr>
<td>Hardness</td>
<td>3</td>
<td>2.3</td>
</tr>
<tr>
<td>Brightness</td>
<td>5</td>
<td>1.4</td>
</tr>
<tr>
<td>Diffuseness</td>
<td>1</td>
<td>1.3</td>
</tr>
<tr>
<td>Naturalness</td>
<td>1</td>
<td>1.3</td>
</tr>
<tr>
<td>Envelopment</td>
<td>1</td>
<td>0.7</td>
</tr>
<tr>
<td>Phasiness</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 3.2: Attribute group, number of occurrences and audibility index obtained for the differences perceived between the images of C-R and L-C-R with cello, bongo and speech sources.

There were a total of 36 stimulus-pairs created for comparison as described in section 2.5. In each attribute test, each subject was asked to compare the 36 stimulus pairs twice, and therefore a total of 72 trial sets were produced. Grading all the 72 trials in one session might have caused experimental errors due to subject fatigue, so the 72 trials were distributed evenly into 3 separate sessions by the type of acoustic condition, each session thus containing 24 trials. In order to avoid such psychological errors as contrast, convergence and anticipation errors [6], the order of presentation for the trials was randomised for each session and for each subject. The orders of sessions and attribute tests were also arranged differently for each subject.

The choice of scale type was influenced by the following considerations. It was thought that using a semantic differential scale with word labels would not be appropriate for this experiment for the following two reasons. Firstly, the potentially nonlinear nature of the scale would not be ideal for parametric statistical analysis. Secondly, the meanings of the labels might be differently interpreted by different subjects. This is likely to be particularly true for an attribute such as source width because it would be difficult for subjects to define the meanings of such labels as ‘much wider’ and ‘slightly wider’ in the same way. With this in mind, using a continuous rating scale was considered to be a more appropriate method since the data would be reliable for parametric statistical analysis due to the linearity of the scale, although the data would need to be normalised before statistical analysis because subjects might use different ranges of the scale. However, using a pure continuous rating scale without any labels, the subjects might have difficulties in maintaining consistency in testing through many trials individually. Therefore, for the scale used in this experiment, 7 point number labels from -30 to 30 were added to a classical continuous rating scale as guidelines for helping subject consistency. The ends of the scale for the locatedness attribute were labelled as ‘more located – less located’, and those for the source width attribute were labelled as ‘wider – narrower’.

The control interface written with MAX-MSP is shown in Figure 4.1. As can be seen, a vertical slider was used for grading, without showing the value to the subjects. The graded value was saved automatically by clicking the ‘next trial’ button. The question presented to the subjects was as shown in the figure, but the order of the
images of C-R and L-C-R presented by the buttons ‘A’ and ‘B’ was randomised for each trial.

![Diagram](image)

Figure 4.1: Layout of the control interface used for the pairwise comparison and grading for source width attribute

Prior to the main grading tests, a few familiarisation trials were provided to the subjects in order to encourage them to use consistent scale ranges and also avoid central tendency errors [6]. 6 representative stimuli comprising the extreme arrays of 1 and 4 combined with 3 sound sources were selected for the familiarisation trials.

4.2. Statistical Analysis

A repeated measure ANOVA (RM ANOVA) was carried out for statistical analysis of the data obtained from the grading test, since all conditions were tested within the same group of subjects. The independent variables were the type of acoustic condition, the type of sound source and the type of microphone array. The dependent variable was the grading of the perceived magnitude of difference between C-R and L-C-R on a scale of -30 to 30.

There were a total of 576 observations, consisting of 16 observations for each of the 36 acoustic-source-array combinations obtained from 8 subjects. Because of the nature of the scale used, it was predicted that each subject would use a different range of the scale. This problem of listener variability in use of the scale might cause inaccurate results from statistical analysis. Therefore, prior to the RM ANOVA test, the original grading data were normalised based on the ITU-R BS.1116 recommendation [15].

Tables 4.1 and 4.2 show the results of RM ANOVA for each attribute test. In the presentation of the results, each independent variable is termed ‘acoustic’, ‘source’ and ‘array’. In order to interpret the results correctly, it was necessary to examine the ‘assumption of sphericity’ (equal variances of the differences between conditions) by using Mauchly’s test of sphericity. Tables 4.3 and 4.4 show the results of Mauchly’s test for each attribute. Insufficient statistic of Mauchly’s test ($p>0.05$) means that the variances of the data for each condition compared are not significantly different, and thus the assumption of sphericity is met. In this case, the ‘sphericity assumed’ significance value should be used as a result of RM ANOVA. However, if Mauchly’s test statistic is significant ($p<0.05$), the assumption of sphericity is violated and one of the corrected significance values should be used instead of the sphericity assumed one.

4.3. Results and Discussions

4.3.1. Source width change

The results of an RM ANOVA test shown in Table 4.1 indicate that microphone array is the most significant factor in source width change ($p=0.000$). The main effect of sound source is also highly significant ($p=0.004$), but the effect size is small ($0.310$) compared to that of microphone array ($0.913$). On the other hand, acoustic condition does not have a significant effect ($p=0.644$). With respect to the interactions between each factor, the largest effect is observed between source and microphone array ($p=0.644$), followed by between acoustic condition and microphone array ($p=0.039$). The acoustic*source interaction is shown to be insignificant ($p=0.714$).

Figure 4.2 shows the mean values and 95% confidence intervals for each microphone array. It can be seen that array 4 has the largest increase of source width when affected by the crosstalk signal, followed by array 3, 2 and 1 in order. Also, there is no overlap of 95% confidence intervals between any pair of arrays, thus causing highly significant differences between all the...
arrays (see Table 4.5). This result suggests that the
effect of interchannel crosstalk on source widening
becomes greater as a more spaced microphone
technique is used, in other words as the ratio of time
difference to intensity difference increased. It also
suggests that this effect can be almost ignored when a
more coincident type of microphone technique is used.
Therefore, this leads to a discussion on the influence of
interchannel time and intensity differences between L
and C. The grounds for this discussion might be the
result of the previous work by the authors [17], showing
that 2 channel stereophonic images were perceived to be
wider compared to the corresponding monophonic
images, and the magnitude of this effect became greater
as the ratio of time difference to intensity difference was
increased. This result might be explained by the effect
of interaural time difference fluctuations (ITD fluctuations)
on the perceived width of a source. Mason
and Rumsey [22] undertook research into interaural
time difference fluctuations (ITD fluctuations) as an
objective measure related to auditory spatial perception
in sound reproduction, and they reported that the
perceived source width increases as the magnitude of
ITD fluctuations becomes greater. In the reproduction
of conventional stereophonic recordings, the amount of
interchannel time difference between each signal can
determine the magnitude of ITD fluctuations. A larger
interchannel time difference will cause a higher degree
of decorrelation between the interaural signals, therefore
a greater magnitude of ITD fluctuations, which also
means a smaller degree of interaural cross correlation
(IACC). In this regard, a spaced microphone technique
will produce a wider phantom image than a coincident
technique due to the difference in the magnitude of ITD
fluctuations. This seems to hold true in the case of the
current experimental conditions. The interchannel time
difference between L and C in the arrays 1 to 4 varies
from 0.5ms to 1.1ms. The longest delay time of the
crosstalk signal L in the array 4 might have caused the
largest change in the magnitude of ITD fluctuations
between the ear input signals of C-R and L-C-R, thus
leading to the largest source width change; whereas the
shortest delay time in the array 1 caused the smallest
change. However, since the microphone techniques
used in this experiment are near-coincident techniques,
it should be noted that it might not only be the
interchannel time difference that contributed to the
significant difference between microphone arrays, but
also the interchannel intensity difference. It was shown
in Table 2.1 that the interchannel intensity difference
between L and C decreases from -20.5dB to -4.6dB as
the microphone array moves from 1 to 4. The decrease
in the intensity difference between L and C means an
increase in the intensity of crosstalk signal. Therefore,
this result seems to suggest that the magnitude of source
width change increases as the intensity of the crosstalk
signal becomes greater. Since the crosstalk signal has a
similar form to a single reflection, the grounds for this
suggestion could be found in concert hall acoustics
research reporting that source width increases as the
intensity of single reflection becomes greater ([23],
[24]). From looking at the mean plots shown in Figure
4.2, it may be possible to ignore the crosstalk effect in
the arrays 1 and 2 since the magnitudes of differences
for those arrays fall within only a 10% region of the
whole scale.

Figure 4.2: Mean value and the associated 95%
confidence intervals of the grade of locatedness
difference between the images of C-R and L-C-R
separated by microphone array

The overall discussion might suggest that a more widely
spaced 3-channel microphone array will tend to give
rise to a greater effect of interchannel crosstalk on
source widening as it will always have greater ITD
fluctuations and greater intensity of the crosstalk signal
due to the nature of near-coincident microphone
technique design requiring a trade-off between
interchannel time and intensity differences. Conversely,
it can also be suggested that in order to minimise the
source width increasing effect of interchannel crosstalk
in the design of 3-channel microphone techniques, one
should pursue a more coincident style of microphone
technique by shortening the delay time and increasing the intensity difference between channels. However, it is not yet known if this source widening effect of interchannel crosstalk can contribute to the improvement of perceived spatial image quality. Experiments studying the spatial impression in concert halls suggest that source widening by reflections is a desirable effect for improving the spatial quality of sound generated in the hall [25]. Regarding the crosstalk signal, L, as a single reflection to C and R could be a way of conceiving of the above question. However, those experiments mostly used much longer delay times of reflections (<80ms) than those of crosstalk signals that might be encountered in general microphone techniques, and therefore preference data obtained in a multichannel stereophonic reproduction are required. The subjective preference of the effect of interchannel crosstalk will be investigated in a future experiment.

Figure 4.3 shows the main effect of sound source on source width change between C-R and L-C-R. It appears that the speech source is outstanding compared to the cello and bongo sources. The multiple comparisons between each sound source indicated in Table 4.6 confirm the significant difference between the speech and the other sources. The cello and bongo are shown to have the same effect (p = 1.000). It seems to be the spectral content that caused the significant difference between speech and the other sources. The effect of spectral content in the signals on source widening has been widely investigated in the research related to concert hall acoustics ([24], [26] and [27]). Barron and Marshall [24] considered that source widening is mainly governed by middle frequencies around 1000 – 2000Hz in the signals while ’envelopment’ is related to low frequencies. Hidaka et al [26] found that the strength factors G below 355Hz (GL : sound pressure level of the sound field at low frequencies) increases source width in a concert hall. The importance of low frequency content in source width was also reported in [27]. In Morimoto and Maekawa’s experiment comparing the source width of noise signals varying in lower cut-off frequency and IACC (interaural cross-correlation), it was found that keeping IACC equal, source width increased as the cut-off frequency decreased below 510Hz and frequency contents around 100 – 200Hz resulted in an especially remarkable increase of source width. These findings seem to provide a reasonable explanation of the current result. As can be seen in Figure 2.6, speech has greater low frequency energies especially at around 100 – 200Hz compared to the other sources, and this seems to be the main factor for the greater source width increase of speech compared to the others, based on the findings of [27].

![Figure 4.2: Mean value and the associated 95% confidence intervals of the grade of locatedness difference between the images of C-R and L-C-R separated by sound source](image)

![Figure 4.4: Mean value and the associated 95% confidence intervals of the grade of locatedness difference between the images of C-R and L-C-R separated by acoustic condition](image)
The main effect of the acoustic condition on source width change is shown in Figure 4.4. Adding multiple reflections and reverberation to an anechoic sound might have increased the source widths for both images of C-R and L-C-R. The insignificant main effect means that the magnitude of the individual increase was similar. This result suggests that the source widening effect of interchannel crosstalk is independent from the acoustic condition of recording space.

The source-array interaction is shown in Figure 4.5. Even though this interaction effect was found to be significant, the order of microphone array in the magnitude of change was the same for all sound sources. Also, considering the estimated effect size is only 0.351, this interaction could possibly be ignored. The acoustic-array interaction was also found to be significant, but again the estimated effect size is shown to be very small (0.135), and the order of microphone array keeps the same regardless of the acoustic condition (see Figure 4.6). Therefore, this interaction could be also ignored. The acoustic-source interaction, which was found to be insignificant is also shown in Figure 4.7.
4.3.2. Locatedness change

Taking an overview of the results of the RM ANOVA test indicated in Table 4.2, ‘microphone array’ shows the most significant main effect to be a locatedness change (the significance value \( p \) is 0.000, and the estimated size of effect is 0.854). The main effect of ‘acoustic condition’ is shown to be significant (\( p = 0.003 \)), but its experimental effect (0.320) is much smaller than that of microphone array. ‘Sound source’ does not have a significant main effect (\( p = 0.637 \)), which means that the magnitude of locatedness change was similar for all sound sources. The largest interaction effect is observed between acoustic and source (\( p = 0.029 \)). The interaction effect of acoustic*array interaction can be judged differently depending on which corrected significance value is used because sphericity is violated (see Table 4.4). That is, the Hyunh-Feldt value (0.043) indicates significance while the Greenhouse-Geisser value (0.052) does not. However, the small partial eta-squared values for acoustic*source (0.162) and acoustic*array (0.172) suggest that the experimental effects of those interactions are relatively minor regardless of the magnitude of the significance value. The source*array interaction is shown to be insignificant (\( p = 0.058 \)).

Figure 4.8 shows the mean value and associated 95% confidence intervals of the grade of locatedness difference between the images of C-R and L-C-R separated by microphone array. The main effect graph for each acoustic condition is shown in Figure 4.9. Even though the graph shows a noticeable decreasing pattern in the magnitude of difference as the microphone array changes from 1 to 4, there is a large overlap between each nearby condition in 95% confidence intervals, which might have led to the relatively small effect size (0.320). The result of a pairwise comparison test shown in Table 4.8 indicates that the only significant difference is between the anechoic and hall conditions (\( p=0.003 \)). In other words, the perceived magnitude of locatedness change was significantly smaller in the hall condition than in the anechoic condition. This is likely to be because the effect of crosstalk was diminished in the hall condition due to the large influence of multiple reflections and reverberation in both images of C-R and L-C-R. This finding might lead to the hypothesis that the effect of crosstalk on locatedness change would become less audible in a more diffused recording space.
The mean values and associated 95% confidence intervals of the normalised data for each sound source are shown in Figure 4.10. As can be seen, all sound sources have small differences in mean values and large overlaps in 95% confidence intervals. Many authors confirmed that a transient sound source is more important for operating the precedence effect than a continuous sound source ([28], [29] and [30]). Therefore, one may presume that the continuous nature of cello source would cause a greater locatedness decrease in the image of L-C-R than the transient nature of bongo source would. However, it should be noted that the characteristics of sound sources used in this experiment are different from those used in the classical studies on the importance of transient sound in localisation. That is, the latter used pure tones, while the former used natural sound sources having complex frequency spectra and waveforms as shown in Figures 2.5 and 2.6. It appears that all sound sources have sufficient transient information to retrigger the precedence effect. For example, the speech source has a fine structure of ongoing transients at every syllable. The cello source also has a continuous musical phrase containing ongoing fluctuations at every note or bow change. Every hit in the bongo source contains a rapid onset transient. Rackerd and Hartmann [30] pointed out that in the case of a complex signal such as noise, the precedence effect could be operated by continuous sounds also. Furthermore, Tobias and Zerlin [31] found that for noise signal, the continuous part became more influential on localisation than the onset transient as the duration of the signal increased. From the above literature, it is assumed that the series of ongoing fluctuations contained in the continuous cello source were strong enough to generate sufficient interaural time differences for retriggering the precedence effect, as well as the strong and rapid transients in the bongo source.

Figure 4.9: Mean value and the associated 95% confidence intervals of the grade of locatedness difference between the images of C-R and L-C-R separated by acoustic condition

Figure 4.10: Mean value and the associated 95% confidence intervals of the grade of locatedness difference between the images of C-R and L-C-R separated by sound source

Figure 4.11 shows the interaction graph between acoustic condition and sound source. There are significant contrasts observed between the anechoic and hall conditions when cello is compared to bongo ($p=0.011$), and when cello is compared to speech ($p=0.028$). In terms of the interaction graph, these contrasts mean that the difference between the cello and the bongo (or speech) in the anechoic condition is significantly bigger than the difference between them in the hall condition. A more detailed interaction can be found in the relationship between each sound source for each acoustic condition. For this investigation, a ‘Paired-Samples T-test’ was performed, and the result summary is shown in Table 4.9. Firstly, in the comparison between sound sources for the anechoic condition, it can be seen that there are significant differences between cello and bongo ($p=0.007$), and between cello and speech ($p=0.048$), although the main effect of sound source is not significant (when acoustic and array are ignored). Bongo and speech do not have a
significant difference. This means that in the anechoic condition, the locatedness decrease is significantly greater in bongo and speech compared with that of cello, but this significance disappears in the room and hall conditions. This result might be explained by the following assumption: the strong transient cues in the bongo and speech signals might have led to a weaker echo suppression effect than the relatively moderate transient cues in the cello signal. This assumption might be supported by Babkoff and Sutton [32]'s finding that if the intensity of a reflection is raised, a perceivable echo appears at a shorter delay time.

Figure 4.12 shows the acoustic-array interaction graph. Array 3 and array 4 have a significant difference when the room and hall conditions are compared. Also, array 2 and array 3 are significantly different when the anechoic and hall conditions are compared. Nevertheless, this effect might be ignored since the order of microphone arrays is the same for all acoustic conditions, and the size of experimental effect is small. This result seems to suggest that the significance of the intensity of crosstalk signal does not change regardless of the acoustic condition of recording space.

The source-array interaction is shown in Figure 4.13. There was no significant interaction between sound source and microphone array. The cello and bongo in array 1 and array 2 are significantly different, but the experimental effect is minor.
4.3.3. Relationship between source width change and locatedness change

Table 4.12 shows the summary of significance values for each attribute. The main effect of microphone array was significant for both locatedness and source width changes. However, the significances of the sound source and acoustic condition effects were found to be opposite in each attribute. That is, the effect of sound source was significant for the source width change, but not for the locatedness change. In contrast, the effect of acoustic condition was significant for the locatedness change, but not for the source width change. For interaction effects also, only the acoustic*source interaction was significant for the locatedness change while it was the only insignificant interaction for the source width change. It seems that the source width and locatedness attributes are often regarded as being negatively correlated. For example, in Berg and Rumsey’s research [33], the ‘source width’ and ‘localisation’, although a different term was used for the same definition, were found to be negatively correlated at a moderate level.

However, the differences found between the locatedness and source width in the current experiment led to a hypothesis that the correlation between those attributes depends on sound source and acoustic condition. Therefore, a set of bivariate correlation tests were carried out. Since the microphone array effects in both attributes have similar tendencies, the level of correlation was expected to be considerable when all the independent variables were included in the test. The result was in fact a moderate negative correlation (-0.670). This means that the ratio of interchannel time and intensity differences affects the changes in both attributes similarly. However, it was also predicted that if only one microphone array was considered, the correlation would be at a low level due to the different main effects of the sound source and acoustic condition. Therefore, individual correlation tests were also performed with each microphone array, and the results confirmed the prediction as can be seen in Table 4.13.

In general this result suggests that with respect to the effect of interchannel crosstalk in a microphone technique, a great source width increase by interchannel crosstalk does not necessarily mean a great locatedness decrease, or vice versa.

<table>
<thead>
<tr>
<th>Main Effect</th>
<th>Interaction Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Array</td>
<td>Source</td>
</tr>
<tr>
<td>Locatedness</td>
<td>0.000</td>
</tr>
<tr>
<td>Source width</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 4.12: Summary of significance values of the main effects and interaction effects for locatedness and source width changes

<table>
<thead>
<tr>
<th>Correlation</th>
<th>Array 1</th>
<th>Array 2</th>
<th>Array 3</th>
<th>Array 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.280</td>
<td>-0.323</td>
<td>-0.169</td>
<td>-0.201</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.13: Correlation value between locatedness change and source width change separated by microphone array
5. SUMMARY AND CONCLUSIONS

A series of subjective experiments were conducted in order to investigate the effect of interchannel crosstalk in multichannel microphone techniques. The independent variables were microphone array type, sound source type, and acoustic condition. The experimental stimuli were created by simulations of multichannel recordings made with the above variables. The experiment employed two processes of attributes selection and grading, and was designed in order that subjects compared the perceptual differences between images with crosstalk and crosstalk-free images. The audible attributes of crosstalk images were first elicited from the subjects, and only the most dominant ones were selected. Then the magnitudes of the selected attributes were graded. The obtained grading data were statistically analysed using the repeated measure ANOVA method. The main findings obtained from the experiments are as follows.

1. The audible attributes of interchannel crosstalk images elicited from the subjects were source width, locatedness, source direction, fullness, source distance, hardness, brightness, diffuseness, naturalness, envelopment, phasiness.

2. The source width and locatedness were found to be the only attributes that were audible more than 'slightly'.

3. In general, the interchannel crosstalk caused increase in source width and decrease in locatedness.

4. Statistically, the magnitudes of both source width increase and locatedness decrease significantly depended on the combination ratio of interchannel time and intensity differences in 3-channel frontal microphone technique. For both attributes, an array employing a greater interchannel time difference (conversely, a greater intensity of crosstalk signal) caused a greater effect.

5. Sound source type was a significant factor for the source width effect but not for the locatedness effect.

6. Acoustic condition had a significant effect on the locatedness decrease, but not on the source width increase.

7. Interactions between microphone array type and sound source type, and between microphone array and acoustic array were significant for the source width effect, but not for the locatedness effect. The experimental effects for these interactions were very small, thus can probably be ignored.

8. Interaction between sound source type and acoustic condition was significant for the locatedness effect, but not for the source width effect. The experimental effect for this interaction was very small, thus can probably be ignored.

9. For each microphone array type, the source width and locatedness effects of interchannel crosstalk had a low correlation.

6. ACKNOWLEDGEMENT

The authors wish to thank the staff and students at the University of Surrey’s Tonmeister course who participated in the experiment, and Yeon-Koo Hong for helping with statistical analysis.

7. REFERENCES


[26] Hidaka, T., Beranek, L. and Okano, T. (1995): ‘Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of


## Tests of Within-Subjects Effects

<table>
<thead>
<tr>
<th>Measure: MEASURE_1</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACOUSTIC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>13.014</td>
<td>2</td>
<td>6.507</td>
<td>.344</td>
<td>.711</td>
<td>.022</td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>13.014</td>
<td>1.450</td>
<td>8.974</td>
<td>.344</td>
<td>.644</td>
<td>.022</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>13.014</td>
<td>1.565</td>
<td>8.317</td>
<td>.344</td>
<td>.660</td>
<td>.022</td>
</tr>
<tr>
<td>Lower-bound</td>
<td>13.014</td>
<td>1.000</td>
<td>13.014</td>
<td>.344</td>
<td>.566</td>
<td>.022</td>
</tr>
<tr>
<td><strong>Error(ACOUSTIC)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>566.708</td>
<td>30</td>
<td>18.890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>566.708</td>
<td>21.751</td>
<td>26.054</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>566.708</td>
<td>23.471</td>
<td>24.146</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-bound</td>
<td>566.708</td>
<td>15.000</td>
<td>37.781</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SOURCE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>495.003</td>
<td>2</td>
<td>247.502</td>
<td>6.733</td>
<td>.004</td>
<td>.310</td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>495.003</td>
<td>1.670</td>
<td>296.437</td>
<td>6.733</td>
<td>.007</td>
<td>.310</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>495.003</td>
<td>1.854</td>
<td>266.956</td>
<td>6.733</td>
<td>.005</td>
<td>.310</td>
</tr>
<tr>
<td>Lower-bound</td>
<td>495.003</td>
<td>1.000</td>
<td>495.003</td>
<td>6.733</td>
<td>.020</td>
<td>.310</td>
</tr>
<tr>
<td><strong>Error(SOURCE)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>1102.719</td>
<td>30</td>
<td>36.757</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>1102.719</td>
<td>25.048</td>
<td>44.025</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>1102.719</td>
<td>27.814</td>
<td>39.646</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-bound</td>
<td>1102.719</td>
<td>15.000</td>
<td>73.515</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ARRAY</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>17141.102</td>
<td>3</td>
<td>5713.701</td>
<td>156.563</td>
<td>.000</td>
<td>.913</td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>17141.102</td>
<td>2.254</td>
<td>7603.759</td>
<td>156.563</td>
<td>.000</td>
<td>.913</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>17141.102</td>
<td>2.673</td>
<td>6412.793</td>
<td>156.563</td>
<td>.000</td>
<td>.913</td>
</tr>
<tr>
<td>Lower-bound</td>
<td>17141.102</td>
<td>1.000</td>
<td>17141.102</td>
<td>156.563</td>
<td>.000</td>
<td>.913</td>
</tr>
<tr>
<td><strong>Error(ARRAY)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>1642.259</td>
<td>45</td>
<td>36.495</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>1642.259</td>
<td>33.814</td>
<td>48.567</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>1642.259</td>
<td>40.094</td>
<td>40.960</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-bound</td>
<td>1642.259</td>
<td>15.000</td>
<td>109.484</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACOUSTIC * SOURCE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>40.007</td>
<td>4</td>
<td>10.002</td>
<td>.531</td>
<td>.714</td>
<td>.034</td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>40.007</td>
<td>2.663</td>
<td>15.022</td>
<td>.531</td>
<td>.643</td>
<td>.034</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>40.007</td>
<td>3.292</td>
<td>12.153</td>
<td>.531</td>
<td>.680</td>
<td>.034</td>
</tr>
<tr>
<td>Lower-bound</td>
<td>40.007</td>
<td>1.000</td>
<td>40.007</td>
<td>.531</td>
<td>.478</td>
<td>.034</td>
</tr>
<tr>
<td><strong>Error(ACOUSTIC * SOURCE)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>1130.771</td>
<td>60</td>
<td>18.846</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>1130.771</td>
<td>39.948</td>
<td>28.306</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>1130.771</td>
<td>49.378</td>
<td>22.900</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-bound</td>
<td>1130.771</td>
<td>15.000</td>
<td>75.385</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACOUSTIC * ARRAY</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>167.944</td>
<td>6</td>
<td>27.991</td>
<td>2.337</td>
<td>.036</td>
<td>.135</td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>167.944</td>
<td>3.602</td>
<td>46.626</td>
<td>2.337</td>
<td>.073</td>
<td>.135</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>167.944</td>
<td>4.881</td>
<td>34.409</td>
<td>2.337</td>
<td>.052</td>
<td>.135</td>
</tr>
<tr>
<td>Lower-bound</td>
<td>167.944</td>
<td>1.000</td>
<td>167.944</td>
<td>2.337</td>
<td>.147</td>
<td>.135</td>
</tr>
<tr>
<td><strong>Error(ACOUSTIC * ARRAY)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>1078.111</td>
<td>90</td>
<td>11.979</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>1078.111</td>
<td>54.030</td>
<td>19.954</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>1078.111</td>
<td>73.212</td>
<td>14.726</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-bound</td>
<td>1078.111</td>
<td>15.000</td>
<td>71.874</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SOURCE * ARRAY</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>630.788</td>
<td>6</td>
<td>105.131</td>
<td>8.097</td>
<td>.000</td>
<td>.351</td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>630.788</td>
<td>2.663</td>
<td>236.855</td>
<td>8.097</td>
<td>.000</td>
<td>.351</td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>630.788</td>
<td>3.292</td>
<td>191.621</td>
<td>8.097</td>
<td>.000</td>
<td>.351</td>
</tr>
<tr>
<td>Lower-bound</td>
<td>630.788</td>
<td>1.000</td>
<td>630.788</td>
<td>8.097</td>
<td>.012</td>
<td>.351</td>
</tr>
<tr>
<td><strong>Error(SOURCE * ARRAY)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphericity Assumed</td>
<td>1168.601</td>
<td>90</td>
<td>12.984</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greenhouse-Geisser</td>
<td>1168.601</td>
<td>39.948</td>
<td>29.253</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Huynh-Feldt</td>
<td>1168.601</td>
<td>49.378</td>
<td>23.687</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower-bound</td>
<td>1168.601</td>
<td>15.000</td>
<td>77.907</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Result table of repeated measure ANOVA test for source width change
Measure: MEASURE_1

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACOUSTIC</td>
<td>Sphericity Assumed</td>
<td>634.292</td>
<td>2</td>
<td>317.146</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>634.292</td>
<td>1.807</td>
<td>351.025</td>
<td>7.063</td>
<td>.004 .320</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>634.292</td>
<td>2.000</td>
<td>317.146</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>634.292</td>
<td>1.000</td>
<td>634.292</td>
<td>7.063</td>
<td>.018 .320</td>
</tr>
<tr>
<td>Error(ACOUSTIC)</td>
<td>Sphericity Assumed</td>
<td>1347.042</td>
<td>30</td>
<td>44.901</td>
<td>1.807</td>
<td>.007 .320</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>1347.042</td>
<td>27.105</td>
<td>49.698</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>1347.042</td>
<td>30.000</td>
<td>44.901</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>1347.042</td>
<td>15.000</td>
<td>89.803</td>
<td>7.063</td>
<td>.018 .320</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>24.385</td>
<td>1.902</td>
<td>12.818</td>
<td>.457</td>
<td>.628 .030</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>24.385</td>
<td>1.000</td>
<td>24.385</td>
<td>.457</td>
<td>.509 .030</td>
</tr>
<tr>
<td>Error(SOURCE)</td>
<td>Sphericity Assumed</td>
<td>799.948</td>
<td>30</td>
<td>26.665</td>
<td>1.807</td>
<td>.105 .320</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>799.948</td>
<td>25.586</td>
<td>31.265</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>799.948</td>
<td>28.537</td>
<td>28.032</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>799.948</td>
<td>15.000</td>
<td>53.330</td>
<td>7.063</td>
<td>.018 .320</td>
</tr>
<tr>
<td>ARRAY</td>
<td>Sphericity Assumed</td>
<td>14067.505</td>
<td>3</td>
<td>4689.168</td>
<td>87.488</td>
<td>.000 .854</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>14067.505</td>
<td>1.325</td>
<td>10619.696</td>
<td>87.488</td>
<td>.000 .854</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>14067.505</td>
<td>1.404</td>
<td>10022.507</td>
<td>87.488</td>
<td>.000 .854</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>14067.505</td>
<td>1.000</td>
<td>14067.505</td>
<td>87.488</td>
<td>.000 .854</td>
</tr>
<tr>
<td>Error ARRAY</td>
<td>Sphericity Assumed</td>
<td>2411.911</td>
<td>45</td>
<td>53.598</td>
<td>1.807</td>
<td>.105 .320</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>2411.911</td>
<td>19.870</td>
<td>121.385</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>2411.911</td>
<td>21.054</td>
<td>114.559</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>2411.911</td>
<td>15.000</td>
<td>160.794</td>
<td>7.063</td>
<td>.018 .320</td>
</tr>
<tr>
<td>ACOUSTIC * SOURCE</td>
<td>Sphericity Assumed</td>
<td>163.104</td>
<td>4</td>
<td>40.776</td>
<td>2.901</td>
<td>.029 .162</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>163.104</td>
<td>2.894</td>
<td>56.367</td>
<td>2.901</td>
<td>.047 .162</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>163.104</td>
<td>3.859</td>
<td>44.576</td>
<td>2.901</td>
<td>.034 .162</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>163.104</td>
<td>1.000</td>
<td>163.104</td>
<td>2.901</td>
<td>.109 .162</td>
</tr>
<tr>
<td>Error(ACOUSTIC*SOURCE)</td>
<td>Sphericity Assumed</td>
<td>843.229</td>
<td>60</td>
<td>14.054</td>
<td>1.807</td>
<td>.105 .320</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>843.229</td>
<td>43.404</td>
<td>19.428</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>843.229</td>
<td>54.885</td>
<td>15.364</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>843.229</td>
<td>15.000</td>
<td>56.215</td>
<td>7.063</td>
<td>.018 .320</td>
</tr>
<tr>
<td>ACOUSTIC * ARRAY</td>
<td>Sphericity Assumed</td>
<td>297.583</td>
<td>6</td>
<td>49.597</td>
<td>3.113</td>
<td>.008 .172</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>297.583</td>
<td>2.230</td>
<td>133.420</td>
<td>3.113</td>
<td>.052 .172</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>297.583</td>
<td>2.638</td>
<td>112.805</td>
<td>3.113</td>
<td>.043 .172</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>297.583</td>
<td>1.000</td>
<td>297.583</td>
<td>3.113</td>
<td>.098 .172</td>
</tr>
<tr>
<td>Error(ACOUSTIC*ARRAY)</td>
<td>Sphericity Assumed</td>
<td>1433.750</td>
<td>90</td>
<td>15.931</td>
<td>1.807</td>
<td>.105 .320</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>1433.750</td>
<td>33.456</td>
<td>42.854</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>1433.750</td>
<td>39.571</td>
<td>36.233</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>1433.750</td>
<td>15.000</td>
<td>95.583</td>
<td>7.063</td>
<td>.018 .320</td>
</tr>
<tr>
<td>SOURCE * ARRAY</td>
<td>Sphericity Assumed</td>
<td>146.698</td>
<td>6</td>
<td>24.450</td>
<td>2.123</td>
<td>.058 .124</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>146.698</td>
<td>3.574</td>
<td>41.042</td>
<td>2.123</td>
<td>.098 .124</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>146.698</td>
<td>4.830</td>
<td>30.371</td>
<td>2.123</td>
<td>.074 .124</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>146.698</td>
<td>1.000</td>
<td>146.698</td>
<td>2.123</td>
<td>.166 .124</td>
</tr>
<tr>
<td>Error(SOURCE*ARRAY)</td>
<td>Sphericity Assumed</td>
<td>1036.302</td>
<td>90</td>
<td>11.514</td>
<td>1.807</td>
<td>.105 .320</td>
</tr>
<tr>
<td></td>
<td>Greenhouse-Geisser</td>
<td>1036.302</td>
<td>53.615</td>
<td>19.329</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Huynh-Feldt</td>
<td>1036.302</td>
<td>72.454</td>
<td>14.303</td>
<td>7.063</td>
<td>.003 .320</td>
</tr>
<tr>
<td></td>
<td>Lower-bound</td>
<td>1036.302</td>
<td>15.000</td>
<td>69.087</td>
<td>7.063</td>
<td>.018 .320</td>
</tr>
</tbody>
</table>

Table 4.2: Result table of repeated measure ANOVA test for locatedness change
### Table 4.3: Mauchly’s test of sphericity for source width change

<table>
<thead>
<tr>
<th>Within Subjects Effect</th>
<th>Mauchly’s W</th>
<th>Approx. Chi-Square</th>
<th>df</th>
<th>Sig.</th>
<th>Epsilon</th>
<th>Greenhouse-Geisser</th>
<th>Huynh-Feldt</th>
<th>Lower-bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACOUSTIC</td>
<td>.621</td>
<td>6.675</td>
<td>2</td>
<td>.036</td>
<td>.725</td>
<td>.782</td>
<td>.500</td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>.802</td>
<td>3.084</td>
<td>2</td>
<td>.214</td>
<td>.835</td>
<td>.927</td>
<td>.500</td>
<td></td>
</tr>
<tr>
<td>ARRAY</td>
<td>.541</td>
<td>8.418</td>
<td>5</td>
<td>.136</td>
<td>.751</td>
<td>.891</td>
<td>.333</td>
<td></td>
</tr>
<tr>
<td>ACOUSTIC * SOURCE</td>
<td>.365</td>
<td>13.505</td>
<td>9</td>
<td>.144</td>
<td>.666</td>
<td>.823</td>
<td>.250</td>
<td></td>
</tr>
<tr>
<td>ACOUSTIC * ARRAY</td>
<td>.095</td>
<td>30.143</td>
<td>20</td>
<td>.075</td>
<td>.600</td>
<td>.813</td>
<td>.167</td>
<td></td>
</tr>
<tr>
<td>SOURCE * ARRAY</td>
<td>.019</td>
<td>50.972</td>
<td>20</td>
<td>.000</td>
<td>.444</td>
<td>.549</td>
<td>.167</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: Mauchly’s test of sphericity for locatedness change

<table>
<thead>
<tr>
<th>Within Subjects Effect</th>
<th>Mauchly’s W</th>
<th>Approx. Chi-Square</th>
<th>df</th>
<th>Sig.</th>
<th>Epsilon</th>
<th>Greenhouse-Geisser</th>
<th>Huynh-Feldt</th>
<th>Lower-bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACOUSTIC</td>
<td>.893</td>
<td>1.582</td>
<td>2</td>
<td>.453</td>
<td>.903</td>
<td>1.000</td>
<td>.500</td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>.827</td>
<td>2.651</td>
<td>2</td>
<td>.266</td>
<td>.853</td>
<td>.951</td>
<td>.500</td>
<td></td>
</tr>
<tr>
<td>ARRAY</td>
<td>.071</td>
<td>36.232</td>
<td>5</td>
<td>.000</td>
<td>.442</td>
<td>.468</td>
<td>.333</td>
<td></td>
</tr>
<tr>
<td>ACOUSTIC * SOURCE</td>
<td>.449</td>
<td>10.736</td>
<td>9</td>
<td>.298</td>
<td>.723</td>
<td>.915</td>
<td>.250</td>
<td></td>
</tr>
<tr>
<td>ACOUSTIC * ARRAY</td>
<td>.022</td>
<td>48.742</td>
<td>20</td>
<td>.000</td>
<td>.372</td>
<td>.440</td>
<td>.167</td>
<td></td>
</tr>
<tr>
<td>SOURCE * ARRAY</td>
<td>.152</td>
<td>24.096</td>
<td>20</td>
<td>.252</td>
<td>.596</td>
<td>.805</td>
<td>.167</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.5: Result of multiple pairwise comparison between each microphone array for source width change

<table>
<thead>
<tr>
<th>(I) ARRAY</th>
<th>(J) ARRAY</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>-1.597</td>
<td>.461</td>
<td>.021</td>
<td>-2.996 - .199</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-8.056</td>
<td>.693</td>
<td>.000</td>
<td>-10.159 - 5.952</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>-13.715</td>
<td>.881</td>
<td>.000</td>
<td>-16.391 - 11.039</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1.597</td>
<td>.461</td>
<td>.021</td>
<td>.199 2.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-6.458</td>
<td>.714</td>
<td>.000</td>
<td>-8.625 - 4.292</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>-12.118</td>
<td>.761</td>
<td>.000</td>
<td>-14.429 - 9.807</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>8.056</td>
<td>.693</td>
<td>.000</td>
<td>5.952 10.159</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6.458</td>
<td>.714</td>
<td>.000</td>
<td>4.292 8.625</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>-5.660</td>
<td>.695</td>
<td>.000</td>
<td>-7.771 - 3.549</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>13.715</td>
<td>.881</td>
<td>.000</td>
<td>11.039 16.391</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>12.118</td>
<td>.761</td>
<td>.000</td>
<td>9.807 14.429</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5.660</td>
<td>.695</td>
<td>.000</td>
<td>3.549 7.771</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Measure: MEASURE_1

<table>
<thead>
<tr>
<th>(I) SOURCE</th>
<th>(J) SOURCE</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cello</td>
<td>Bongo</td>
<td>.125</td>
<td>.463</td>
<td>1.000</td>
<td>Lower Bound: -1.122 Upper Bound: 1.372</td>
</tr>
<tr>
<td></td>
<td>Speech</td>
<td>-1.901</td>
<td>.699</td>
<td>.047</td>
<td>Lower Bound: -3.784 Upper Bound: -.016</td>
</tr>
<tr>
<td>Bongo</td>
<td>Cello</td>
<td>-1.125</td>
<td>.463</td>
<td>1.000</td>
<td>Lower Bound: -1.372 Upper Bound: 1.122</td>
</tr>
<tr>
<td></td>
<td>Speech</td>
<td>-2.026</td>
<td>.667</td>
<td>.025</td>
<td>Lower Bound: -3.824 Upper Bound: -.228</td>
</tr>
<tr>
<td>Speech</td>
<td>Cello</td>
<td>1.901</td>
<td>.699</td>
<td>.047</td>
<td>Lower Bound: .018 Upper Bound: 3.784</td>
</tr>
<tr>
<td></td>
<td>Bongo</td>
<td>2.026</td>
<td>.667</td>
<td>.025</td>
<td>Lower Bound: .228 Upper Bound: 3.824</td>
</tr>
</tbody>
</table>

Table 4.6: Result of multiple pairwise comparison between each sound source for source width change

Measure: MEASURE_1

<table>
<thead>
<tr>
<th>(I) ARRAY</th>
<th>(J) ARRAY</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1.625</td>
<td>.287</td>
<td>.000</td>
<td>Lower Bound: -2.495 Upper Bound: 2.495</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>7.625</td>
<td>.766</td>
<td>.000</td>
<td>Lower Bound: 5.300 Upper Bound: 9.950</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>12.424</td>
<td>1.251</td>
<td>.000</td>
<td>Lower Bound: 8.626 Upper Bound: 16.221</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>-1.625</td>
<td>.287</td>
<td>.000</td>
<td>Lower Bound: -4.206 Upper Bound: -.898</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>6.000</td>
<td>.677</td>
<td>.000</td>
<td>Lower Bound: 3.944 Upper Bound: 8.056</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>10.799</td>
<td>1.117</td>
<td>.000</td>
<td>Lower Bound: 7.409 Upper Bound: 14.189</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>-7.625</td>
<td>.766</td>
<td>.000</td>
<td>Lower Bound: -9.950 Upper Bound: -5.300</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>-6.000</td>
<td>.677</td>
<td>.000</td>
<td>Lower Bound: -8.056 Upper Bound: -3.944</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4.799</td>
<td>.727</td>
<td>.000</td>
<td>Lower Bound: 2.590 Upper Bound: 7.007</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>12.424</td>
<td>1.251</td>
<td>.000</td>
<td>Lower Bound: -6.621 Upper Bound: -8.626</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>-10.799</td>
<td>1.117</td>
<td>.000</td>
<td>Lower Bound: -14.189 Upper Bound: -7.409</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>-4.799</td>
<td>.727</td>
<td>.000</td>
<td>Lower Bound: -7.007 Upper Bound: -2.590</td>
</tr>
</tbody>
</table>

Table 4.7: Result of multiple pairwise comparison between each microphone array for locatedness change

Measure: MEASURE_1

<table>
<thead>
<tr>
<th>(I) ACOUSTIC</th>
<th>(J) ACOUSTIC</th>
<th>Mean Difference (I-J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval for Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anechoic</td>
<td>Room</td>
<td>-.1542</td>
<td>.787</td>
<td>.207</td>
<td>Lower Bound: -3.662 Upper Bound: .579</td>
</tr>
<tr>
<td></td>
<td>Hall</td>
<td>-2.552</td>
<td>.614</td>
<td>.003</td>
<td>Lower Bound: -4.206 Upper Bound: -.898</td>
</tr>
<tr>
<td>Room</td>
<td>Anechoic</td>
<td>1.542</td>
<td>.787</td>
<td>.207</td>
<td>Lower Bound: .759 Upper Bound: 3.662</td>
</tr>
<tr>
<td></td>
<td>Hall</td>
<td>-1.010</td>
<td>.638</td>
<td>.402</td>
<td>Lower Bound: -2.728 Upper Bound: 7.07</td>
</tr>
<tr>
<td>Hall</td>
<td>Anechoic</td>
<td>2.552</td>
<td>.614</td>
<td>.003</td>
<td>Lower Bound: .898 Upper Bound: 4.206</td>
</tr>
<tr>
<td></td>
<td>Room</td>
<td>1.010</td>
<td>.638</td>
<td>.402</td>
<td>Lower Bound: -.707 Upper Bound: 2.728</td>
</tr>
</tbody>
</table>

Table 4.8: Result of multiple pairwise comparison between each acoustic condition for locatedness change
<table>
<thead>
<tr>
<th>Pair</th>
<th>Conditions</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>anechoic+cello - anechoic+bongo</td>
<td>.016</td>
</tr>
<tr>
<td>2</td>
<td>anechoic+cello - anechoic+speech</td>
<td>.048</td>
</tr>
<tr>
<td>3</td>
<td>anechoic+bongo - anechoic+speech</td>
<td>.678</td>
</tr>
<tr>
<td>4</td>
<td>room+cello - room+bongo</td>
<td>.611</td>
</tr>
<tr>
<td>5</td>
<td>room+cello - room+speech</td>
<td>.145</td>
</tr>
<tr>
<td>6</td>
<td>room+bongo - room+speech</td>
<td>.250</td>
</tr>
<tr>
<td>7</td>
<td>hall+cello - hall+bongo</td>
<td>.241</td>
</tr>
<tr>
<td>8</td>
<td>hall+cello - hall+speech</td>
<td>.125</td>
</tr>
<tr>
<td>9</td>
<td>hall+bongo - hall+speech</td>
<td>.686</td>
</tr>
<tr>
<td>10</td>
<td>anechoic+cello - room+cello</td>
<td>.192</td>
</tr>
<tr>
<td>11</td>
<td>anechoic+cello - hall+cello</td>
<td>.132</td>
</tr>
<tr>
<td>12</td>
<td>room+cello - hall+cello</td>
<td>.892</td>
</tr>
<tr>
<td>13</td>
<td>anechoic+bongo - room+bongo</td>
<td>.000</td>
</tr>
<tr>
<td>14</td>
<td>anechoic+bongo - hall+bongo</td>
<td>.000</td>
</tr>
<tr>
<td>15</td>
<td>room+bongo - hall+bongo</td>
<td>.131</td>
</tr>
<tr>
<td>16</td>
<td>anechoic+speech - room+speech</td>
<td>.171</td>
</tr>
<tr>
<td>17</td>
<td>anechoic+speech - hall+speech</td>
<td>.000</td>
</tr>
<tr>
<td>18</td>
<td>room+speech - hall+speech</td>
<td>.010</td>
</tr>
</tbody>
</table>

Table 4.9: Result of paired T-test for acoustic condition and sound source